

DOI: 10.2436/20.1501.02.85

Cent cinquanta anys després de *L'origen de les espècies*, de Darwin  
(Arcadi Navarro i Carmen Segarra, ed.)*Treballs de la SCB. Vol. 60 (2009) 195-204*

## COMPLINT UNA INTUÏCIÓ DE DARWIN: GENÈTICA HUMANA I LLENGÜES

DAVID COMAS, ELENA BOSCH I FRANCESC CALAFELL

*Institut de Biologia Evolutiva (UPF-CSIC), Departament de Ciències Experimentals i de la Salut,  
Universitat Pompeu Fabra i CIBER de Salut Pública.*

Adreça per a la correspondència: Francesc Calafell. Unitat de Biologia Evolutiva,  
Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra.  
Doctor Aiguader, 88. 08003 Barcelona. Adreça electrònica: [francesc.calafell@upf.edu](mailto:francesc.calafell@upf.edu).

### RESUM

Les similituds entre la manera com s'hereten i evolucionen els gens i les llengües han promogut l'interès en l'anàlisi conjunta d'ambdós camps. La correlació entre gens i llengües va ser demostrada per primer cop per Cavalli-Sforza i col·laboradors en comparar un arbre construït amb marcadors genètics clàssics amb un arbre lingüístic de les llengües. S'han formulat diverses crítiques a aquest mètode, i se n'han descrit nombroses excepcions. S'ha demostrat que el factor més important que determina la correlació de gens i llengües és la geografia. Nogensmenys, la informació proporcionada per la genètica i la lingüística, ensems amb la proporcionada per altres disciplines, ens permetrà reconstruir la història de la humanitat.

**Paraules clau:** diversitat genètica, diversitat lingüística, correlació gens-llengües, coevolució gens-cultura, Cavalli-Sforza.

### FULFILLING ONE OF DARWIN'S INTUITIONS: HUMAN GENETICS AND LANGUAGES

### SUMMARY

The similarities between the mode of inheritance and evolution of genes and languages have fostered interest in the joint analysis of both disciplines. The correlation between genes and languages was firstly demonstrated by Cavalli-Sforza and collaborators comparing a tree build from 'classical' genetic markers to a linguistic tree of languages. Several criticisms to this correlation have been raised and a large number of exceptions have been described. It has been shown that the most plausible factor that influences in the cor-

relation of genes and languages is geography. However, the information provided by genetics and linguistics, as well as the one provided by other disciplines, will allow us to reconstruct the history of humankind.

**Key words:** genetic diversity, linguistic diversity, gene-languages correlation, gene-culture coevolution, Cavalli-Sforza.

*If we possessed a perfect pedigree of mankind, a genealogical arrangement of the races of man would afford the best classification of the various languages now spoken throughout the world.*

*On the origin of species*, capítol XIII.

## INTRODUCCIÓ

El lligam entre l'evolució de l'espècie humana i les llengües ja fou assenyalat per Charles Darwin a *L'origen de les espècies* (1859), on suggerí que la reconstrucció de l'arbre evolutiu humà permetria refinar la classificació de les llengües humanes. Les dades provinents de la paleoantropologia i el desenvolupament de l'antropologia molecular ens han permès reconstruir l'arbre evolutiu dels humans actuals de manera molt acurada. En canvi, la reconstrucció de l'arbre de les llengües, especialment en les seves branques més profundes, roman poc detallat i controvertit, a causa de l'evolució més ràpida de les llengües i, com veurem més endavant, dels diferents processos que fan que l'evolució de les llengües s'aparti de la manera arborescent estricta que presenta l'evolució genètica.

Sens dubte, la parla és un dels trets que defineix la humanitat. No és l'objectiu d'aquest article descriure les característiques anatòmiques, fisiològiques, neurològiques i, en definitiva, genètiques, que fan possible la parla humana, ja que ens centrarem no en allò que gairebé tots els humans compartim (és a dir, la capacitat de parlar), sinó en allò que és diferent entre

humans (val a dir, quina llengua parlem). Tanmateix, resumirem algunes de les troballes recents més importants sobre l'evolució de la parla. S'ha trobat que el gen *FOXP2* està implicat en la capacitat lingüística (Lai *et al.*, 2001), i les anàlisis moleculars assenyalen que aquest gen mostra adaptació recent en el llinatge humà (Enard *et al.*, 2002). Aquestes anàlisis han mostrat que aquest gen ha sofert adaptació recent en els humans, comparant-ho amb altres primats, cosa que suggereix que l'aparició i expansió dels humans moderns, fa dos-cents mil anys, podria haver estat impulsada per l'adquisició d'una parla sofisticada. Tanmateix, els resultats de l'anàlisi del DNA de dos neandertals (Krause *et al.*, 2007) mostren que els al·lels derivats dels humans ja eren presents en els neandertals. Per tant, l'aparició d'aquests canvis en el gen *FOXP2* és anterior a la separació dels llinatges que van portar als neandertals i als humans moderns.

La parla humana és una característica universal: tots els grups humans han desenvolupat una llengua sofisticada, encara que moltes llengües no hagin disposat d'un sistema d'escriptura fins a temps molt recents. Gairebé tots els humans tenim la mateixa aptitud biològica per aprendre una llengua. En canvi, no està genèticament determinat quina llengua concreta es parla, i cada individu aprèn el vocabulari i la gramàtica d'una llengua específica segons en quin context cultural neix. Això és particularment evident en les adopcions internacionals: els nens adoptats aprenen fàcilment la llengua dels seus nous pares. Les llengües són part de la cultura humana i, per

tant, es transmeten de pares a fills, de generació en generació, de la mateixa manera que es transmeten altres trets culturals com la religió, la tecnologia o l'ètica. Aquest fet té implicacions dràstiques en la correlació entre gens i llengües. A causa del seu mode cultural d'herència i transmissió, les llengües es poden influir, modificar o reemplaçar fàcilment per altres llengües, i aquests canvis poden passar ràpidament, en menys d'una generació.

La reconstrucció de la història evolutiva comuna de dos individus, grups, poblacions o espècies es pot dur a terme utilitzant les eines de la genètica evolutiva. El principi bàsic, en la seva forma més simple, és el següent: com més semblants genèticament són dues entitats, més recent és el seu avantpassat comú. Aquest principi bàsic implica que les diferències genètiques s'acumulen al llarg del temps i, per tant, com més temps ha passat des de la separació de dues entitats, més grans seran les diferències genètiques entre aquestes. Paralelament, la similitud lingüística que observem actualment entre les llengües apunta al seu origen comú en una llengua ancestral o protollengua. A més dels paral·lelismes entre genètica i lingüística, la substitució lingüística pot ser més ràpida que el canvi genètic, ja que la transmissió genètica és vertical (és a dir, de pares a fills), mentre que la transmissió lingüística pot ser vertical o horitzontal, és a dir, d'un individu a un altre, emparentats o no, dins d'una població. A més, alguns processos, tant en genètica com en lingüística, poden distorsionar les diferències entre poblacions, i poden fer que poblacions allunyades apareguin molt més properes entre si del que són realment. En genètica, l'evolució convergent, l'homoplàsia, l'adaptació o la mescla poden disminuir o esborrar les diferències genètiques acumulades al llarg del temps des de la separació de dues poblacions. Per exemple, poblacions com els

centreafricans i els aborígens australians, que es van separar fa més de cinquanta mil anys, poden presentar adaptacions genètiques a la radiació solar semblants en gens relacionats amb el color de la pell, malgrat els seus remots orígens comuns. De la mateixa manera, processos com els manlleus lingüístics o la substitució d'idiomes poden mostrar similituds lingüístiques degudes a esdeveniments recents no relacionats amb la separació de les llengües.

Ambdues disciplines, la genètica i la lingüística, utilitzen eines semblants per a la reconstrucció de les històries dels gens i les llengües, respectivament. Les distàncies genètiques i lingüístiques es poden calcular de diferents maneres en ambdues disciplines i es poden representar després com a arbres. Les distàncies genètiques es poden calcular a partir de diferències en les freqüències al·lèliques o de diferències en seqüències de DNA. De manera semblant, els lingüistes poden usar llistes de mots comuns, com la llista de Swadesh (vegeu la taula 1) per quantificar les diferències entre idiomes. Encara que els arbres produeixen una representació visual atractiva de les dades, imposen un model bifurcant a una matriu de distàncies que pot no presentar aquesta estructura, especialment entre entitats estretament relacionades. A més, els arbres no poden representar processos com els manlleus lingüístics o la mescla entre poblacions, de manera que l'aplicació dels arbres és limitada i la seva interpretació pot ser esbiaixada. De totes maneres, si el propòsit de l'anàlisi és determinar si existeix una correlació entre gens i llengües, no cal utilitzar arbres: una correlació entre les matrius de distàncies genètiques i lingüístiques es pot calcular directament a partir de les matrius mateixes, mitjançant el test de Mantel. A més, en aquest tipus d'anàlisi es pot descomptar l'efecte de variables que poden confondre, com la distància

TAULA 1. Un subconjunt de vint-i-cinc paraules de la llista de Swadesh, en català, francès, italià, llatí, anglès, holandès, alemany i suec.

Català	Francès	Italià	Llatí	Anglès	Holandès	Alemany	Suec
jo	je	io	ego	I	ik	ich	jäg
tu, vostè, vós (formal)	tu, vous (formal)	tu, lei (formal)	tu	you (s.), thou	jij, je, U (formal)	du, Sie (formal)	du
ell	il	lui, egli	is, ea	he	hij	er	han
nosaltres	nous	noi	nos	we	wij, we	wir	vi
vosaltres, vos- tès (formal)	vous	voi	vos	you	jullie	ihr, Sie (formal)	ni
ells, elles	ils, elles	loro, essi	ii, eae	they	zij, ze	sie	de
aquest	ceci	questo	hic, is	this	Deze, dit	dieses	dethär
aquell	cela	quello	ille	that	die, dat	jenes	detdär
aquí	ici	qui, qua	hic	here	hier	hier	här
allà	là	là	ibi	there	daar	dort	där
qui	qui	chi	quis	who	wie	wer	vem
què	quoi	che	quid	what	wat	was	vad
on	où	dove	ubi	where	waar	wo	var
quan	quand	quando	quando	when	wanneer	wann	när
com	comment	come	quam, quomodo	how	hoe	wie	hur
no	ne...pas	non	non	not	niet	nicht	inte, ej
tot	tout	tutto	omnis	all	al, alle	alle	alla
molts	plusieurs	molti	multi	many	veel	viele	många
alguns, uns	quelques	alcuni	aliqui, ali- quot	some	enkele, sommige	einige	några, vissa
poc	peu	pochi	pauci	few	wenig	wenige	få
altres	autre	altro	alter, alius	other	ander	andere	annan
un	un	uno	unus	one	een	eins	ett
dos	deux	due	duo	two	twee	zwei	två
tres	trois	tre	tres	three	drie	drei	tre
quatre	quatre	quattro	quattuor	four	vier	vier	fyra

geogràfica; podeu veure exemples d'aquesta aproximació a Rosser *et al.* (2000), Rubicz *et al.* (2002), Karafet *et al.* (2002), Ayub *et al.* (2003) i Wood *et al.* (2005), entre altres.

## CORRELACIÓ ENTRE GENS I LLENGÜES

Les expansions demogràfiques humanes, impulsades per innovacions tecnològiques o culturals, poden implicar la dispersió dels individus i, per tant, l'expansió dels

seus gens i les seves llengües a nous territoris. La colonització d'un nou territori com a conseqüència d'una expansió demogràfica implica inicialment una similitud genètica entre els colonitzadors, els quals poden divergir posteriorment. Per tant, podem inferir que les poblacions extenses que comparteixen una llengua o un grup de llengües estretament emparentades poden ser el resultat d'expansions de poblacions en què encara no ha passat prou temps per a la divergència lingüística. De la mateixa manera, es pot inferir que les poblacions que ocu-

pen un cert territori i que parlen llengües diverses de la mateixa família lingüística presenten un origen comú més antic (vegeu la figura 1). Alguns exemples ben coneguts de l'expansió de famílies lingüístiques són els casos de l'indoeuropeu, el bantu o l'austroasiàtic. Aquestes expansions lingüístiques es correlacionen gairebé perfectament amb expansions genètiques, i les poblacions que parlen actualment aquestes llengües també estan estretament emparentades genèticament. Moltes d'aquestes expansions van ser provocades per la dispersió dels productors d'aliments i de les seves llengües cap als territoris ocupats per poblacions de caçadors-recol·lectors. Com a resultat d'aquest procés, la distribució d'aquestes famílies lingüístiques (i gens) és àmplia. La distribució de les famílies indoeuropea, sinotibetana, bantu i austronèsia s'ha associat amb les expansions de l'agricultura des de l'Orient Mitjà, el nord de la Xina, la frontera entre Nigèria i el Camerun, i el sud de la Xina, respectivament. Després d'aquestes expansions, les llengües (i els gens) es poden haver començat de diferenciar i, per tant, el que observem avui és un grup de distribucions contigües de llengües i gens relacionats. Tanmateix, l'argument de l'expansió conjunta dels gens i les llengües és tan vàlid com una de les seves principals premisses, val a dir, la difusió de l'agricultura, en què la innovació cultural s'escampa a través de la migració de les poblacions implica des que, per tant, escampen també els seus gens. En canvi, altres processos com l'aculturació, que implica la difusió d'innovacions culturals sense el reemplaçament de les poblacions, poden comportar una substitució cultural (incloent-hi probablement la lingüística) sense un recanvi de gens.

Les expansions neolítiques constitueixen una categoria de moviments poblacionals que poden haver escampat simultàniament gens i llengües. Renfrew (1994) va

proposar un marc teòric general en què la localització geogràfica actual de les principals famílies lingüístiques està lligada a un o més d'un petit conjunt de moviments (pre)històrics de població. Breument: *a*) algunes famílies lingüístiques (khoisanid, nilosahariana, caucàsica, australiana, indopacífica i ameríndia; vegeu la figura 1) deuen les seves localitzacions geogràfiques només a la dispersió inicial dels humans anatòmicament moderns en el paleolític superior; *b*) les dispersions dels humans neolítics haurien escampat les famílies nigerokurdofànica, afroasiàtica, indoeuropea, elamodravídica, sinotibetana i austronèsiana; *c*) un clima més suau a les latituds boreals podria haver portat la colonització de les regions més septentrionals i l'expansió de les famílies uràlica, txuktxikamtxatka, nadené i eskimoaleutiana i *d*) moviments posteriors sota un model de dominància d'elit, és a dir, en què una minoria ben organitzada pot conquerir un territori i imposar una cultura i una llengua; aquest fou el cas de l'expansió de la família altaica des d'un bressol relativament reduït a l'Àsia Central, el de l'indoeuropeu al subcontinent indi, o el de la introducció de la família indoeuropea mateixa a gran part del món en la colonització europea postcolombiana.

L'expansió conjunta de l'agricultura i de les poblacions (o, de fet, de qualsevol dels processos *a* a *c* anteriors) pot crear una correlació entre les distàncies o els arbres evolutius de gens o llengües. Cavalli-Sforza (1997) i Cavalli-Sforza *et al.* (1998) foren els primers a demostrar aquesta correlació entre gens i llengües. Van recollir freqüències al·lèliques de cent vint polimorfismes genètics clàssics (és a dir, d'aquells que es detecten en els productes dels gens i no pas en els gens mateixos, com els grups sanguinis i altres polimorfismes proteics) en quaranta-dues poblacions d'arreu del món, van calcular-ne les distàncies genètiques i les

van representar en forma d'arbre. A continuació, van comparar aquest arbre genètic amb un de lingüístic (Ruhlen, 1991), i es va mostrar una alta correlació entre ambdós arbres (vegeu la figura 2). Van trobar algunes excepcions a aquesta correlació global, com algunes poblacions pertanyents a la mateixa família lingüística però genèticament molt divergents o, al contrari, poblacions genèticament properes que parlaven llengües de famílies diferents. Tanmateix, diversos autors van criticar les dades i els mètodes usats per Cavalli-Sforza i els seus col·laboradors (Bateman *et al.*, 1990). Alguns dels punts criticats foren la definició arbitrària de les poblacions, la classificació lingüística utilitzada, l'escassa mida d'algunes mostres i el mètode usat per a la reconstrucció de l'arbre.

Malgrat les nombroses excepcions descrites, la correlació entre gens i llengües continua sent bona. Però aquesta correlació es podria atribuir a factors diferents d'un lligam directe entre gens i llengües, com podria ser la correlació d'ambdues entitats

amb una tercera variable. El factor més versemblant que podria influir en la correlació entre gens i llengües és la geografia: si els gens estan correlacionats amb la geografia perquè els moviments humans, les expansions i les migracions estan constretes per les característiques de la geografia, i si les llengües es difonen de manera limitada per les barreres geogràfiques mateixes, llavors els gens i les llengües podrien estar correlacionats entre si perquè ambdós estarien correlacionats amb la geografia. Algunes dades genètiques han mostrat una alta correlació amb la geografia però no amb les llengües. Per exemple, Rosser *et al.* (2000) demostren que la diversitat genètica actual en el cromosoma Y a Europa mostra un patró clinal altament correlacionat amb la geografia però no amb les llengües, cosa que suggeria una expansió des de l'Orient Mitjà amb heterogeneïtats lingüístiques.

S'ha trobat una correlació diferencial entre gens i llengües segons el marcador genètic utilitzat, cosa que presenta implicacions potents. S'ha vist que la correlació entre

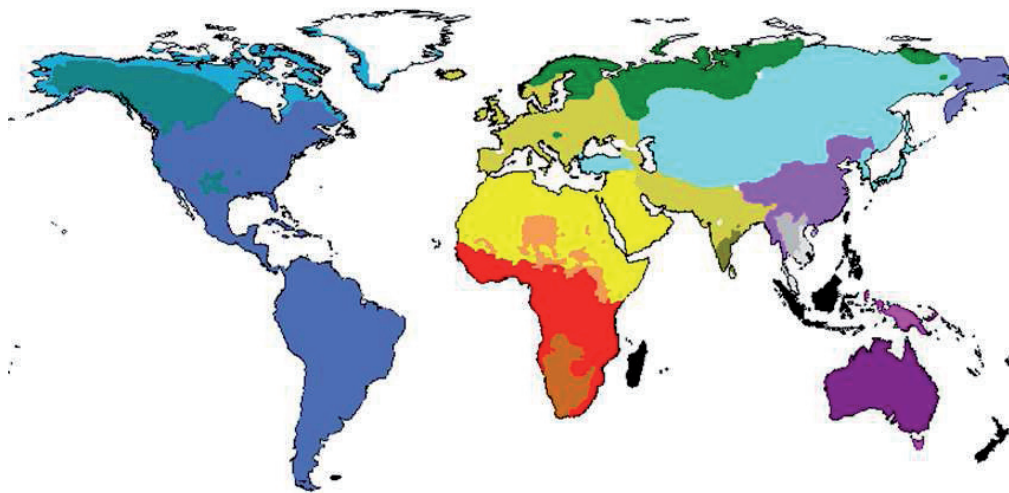


FIGURA 1. Mapa de les principals famílies lingüístiques, adaptat de Ruhlen (1991). Les famílies lingüístiques estan aco-  
lorides com a la figura 2. Les àrees blanques dins dels continents representen zones deshabitades o isolats lingüístics.



la diversitat genètica trobada en el cromosoma Y i la diversitat lingüística és més alta que no pas la correlació entre DNA mitocondrial i llengües (Poloni *et al.*, 1997; Pérez-Lezaun *et al.*, 1999; Bosch *et al.*, 2006). Aquest fet es podria explicar per les diferents taxes de migració masculina i femenina. S'ha postulat (Seielstad *et al.*, 1998) que la taxa de migració femenina és més alta que la masculina, cosa que podria explicar per què els llinatges masculins són espacialment més heterogenis que els llinatges femenins. Encara que la llengua pròpia d'un

individu se sol anomenar la llengua *materna*, la millor correlació de la diversitat lingüística amb la diversitat genètica transmesa pels homes suggereix que *llengua paterna* seria una expressió més adequada.

## PAISATGES GENÈTICS I LINGÜÍSTICS

La composició genètica de les poblacions, és a dir, les seves freqüències al·lèliques o haplotípiques en els locus rellevants, can-

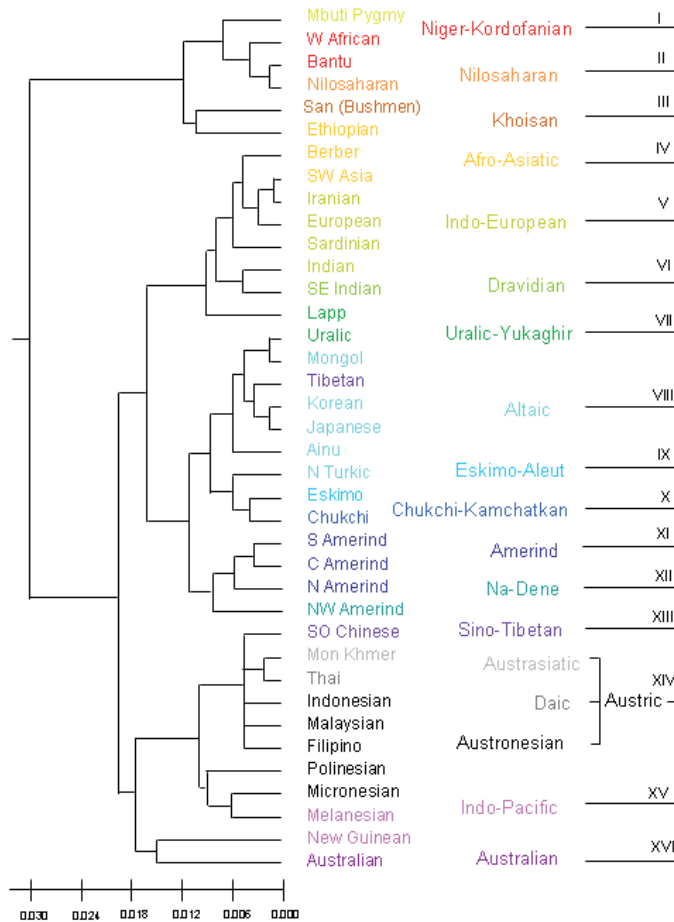


FIGURA 2. Arbres que relliquen la diversitat genètica (esquerra) i lingüística (dreta).

via en l'espai. Òbviament, un determinant principal de la taxa de canvi són els obstacles geogràfics al moviment de les poblacions (i, per tant, al flux genètic) com els mars i les serralades. Tanmateix, les diferències lingüístiques poden tenir també un paper en la creació del paisatge genètic, tot dificultant l'aparellament de persones que parlin llengües diferents i reforçant l'endogàmia d'un grup. Barbuani i Sokal (1990) analitzaren la taxa de canvi en l'espai de les freqüències al·lèliques a Europa, determinaren on era més pronunciat el canvi genètic, i trobaren que aquestes fronteres genètiques coincidien amb fronteres lingüístiques amb una probabilitat més gran de l'esperada si ambdós tipus de fronteres fossin independents entre si. Sovint, les barreres lingüístiques i genètiques coincidien en llocs on no hi havia barreres geogràfiques, cosa que implicava que les llengües havien restringit realment el flux genètic entre poblacions a Europa. En especial, es va trobar una clara frontera genètica al voltant dels bascos (Bertranpetit i Calafell, 1994), que no estan separats de la resta de la península Ibèrica per cap barrera geogràfica particularment difícil. Una aproximació similar es va aplicar a les poblacions d'Itàlia (Zei *et al.*, 1993; en aquest cas, la diferenciació genètica es va estimar a partir dels cognoms, que es poden modelar com un sol locus del cromosoma Y amb molts allels), la Gran Bretanya (Falsetti i Sokal, 1993) i el Japó (Sokal i Thomson, 1998).

## EXCEPCIONS EN LA CORRELACIÓ ENTRE GENS I LLENGÜES

La correlació entre gens i llengües ja es dona per esperada i es pot considerar, fent un símil estadístic, com una hipòtesi nul·la; així, les excepcions a aquesta norma són més interessants i ens donen pistes sobre

les històries de moltes poblacions locals. Les excepcions en els estudis seminals de Cavalli-Sforza i col·laboradors foren atribuïdes als fenòmens de substitució lingüística o poblacional. Alguns exemples ben coneguts de substitució lingüística són els pigmeus, que parlen llengües del grup bantu o de la família nilosahariana, mentre que romanen clarament diferenciats genèticament dels parlants no pigmeus d'aquests grups lingüístics. Els exemples de substitució lingüística són abundants en les antigues colònies europees, on molts grups nadius van aconseguir retenir la seva diferenciació genètica, mentre les seves llengües eren reemplaçades per la de la metròpoli colonial. Aquests són exemples del model de dominància d'elit en el marc general de Renfrew. En aquest cas, un grup limitat d'individus pren el control polític, social o religiós d'una població general i imposa una llengua. En conseqüència, la substitució lingüística pot tenir lloc sense la substitució genètica de la població. Aquest procés és possible només amb poblacions molt estructurades amb una jerarquia social, en què el reemplaçament d'un petit nombre d'individus del cim de la piràmide social pot afectar la població general. Un exemple addicional d'aquest model de dominància d'elit el constitueix la llengua túrquica parlada actualment a Turquia. En el segle *xi*, tribus túrquiques provinents d'Àsia Central van imposar la seva llengua a les poblacions anatòliques, i van substituir el grec i altres llengües indoeuropees, algunes de les quals van desaparèixer. Actualment, la distribució geogràfica de les llengües indoeuropees és discontinua a causa de la substitució a Anatòlia, però la composició genètica dels turcs roman més propera a les poblacions de l'Orient Mitjà i Europa que no pas a les de l'Àsia Central (Calafell *et al.*, 1996; Comas *et al.*, 1996; Cinnioglu *et al.*, 2004). De manera semblant, els



descendents culturals dels romans als Balcans, és a dir, els romanesos i els aromuns, ambdós parlants de llengües romàniques, són genèticament més semblants als balcànics parlants d'altres llengües (principalment eslaves) que no pas als italians (Bosch *et al.*, 2006).

L'anàlisi conjunta dels gens i les llengües es troba en la intersecció del coneixement fornit per diverses disciplines diferents de la genètica i la lingüística, com l'arqueologia, les ciències cognitives i la paleontologia, entre altres. La conseqüència final de treure l'entrellat de les relacions entre gens i llengües ens portarà no solament a la reconstrucció de les migracions i dels intercanvis culturals entre poblacions, sinó també a la singular història evolutiva de la nostra espècie, tal com Darwin havia intuït.

## BIBLIOGRAFIA

- AYUB, Q.; MANSOOR, A.; ISMAIL, M. [et al.] (2003). «Reconstruction of human evolutionary tree using polymorphic autosomal microsatellites». *Am. J. Phys. Anthropol.*, 122: 259-268.
- BARBUJANI, G.; SOKAL, R. R. (1990). «Zones of sharp genetic change in Europe are also linguistic boundaries». *Proc. Nat. Acad. Sci. USA*, 87: 1816-1819.
- BATEMAN, R.; GODDARD, I.; O'GRADY, R. [et al.] (1990). «Speaking of forked tongues: the feasibility of reconciling human phylogeny and the history of language». *Curr. Anthropol.*, 31: 1-13.
- BOSCH, E.; CALAFELL, F.; GONZÁLEZ-NEIRA, A. [et al.] (2006). «Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns». *Ann. Hum. Gen.*, 70: 459-487.
- CALAFELL, F.; BERTRANPETIT, J. (1994). «Principal component analysis of gene frequencies and the origin of the Basques». *Am. J. Phys. Anthropol.*, 93: 201-215.
- CALAFELL, F.; UNDERHILL, P.; TOLUN, A.; ANGELICHEVA, D.; KALAYDJIEVA, L. (1996). «From Asia to Europe: mitochondrial DNA sequence variability in Bulgarians and Turks». *Ann. Hum. Gen.*, 60: 35-49.
- CAVALLI-SFORZA, L. (1997). «Genes, peoples, and languages». *Proc. Nat. Acad. Sci. USA*, 94: 7719-7724.
- CAVALLI-SFORZA, L.; PIAZZA, A.; MENOZZI, P.; MOUNTAIN, J. (1998). «Reconstruction of human evolution: bringing together genetic, archaeological, and linguistic data». *Proc. Nat. Acad. Sci. USA*, 85: 6002-6006.
- CINIOGLU, C.; KING, R.; KIVISILD, T. [et al.] (2004). «Excavating Y-chromosome haplotype strata in Anatolia». *Hum. Gen.*, 114: 127-148.
- COMAS, D.; CALAFELL, F.; MATEU, E.; PÉREZ-LEZAUN, A.; BERTRANPETIT, J. (1996). «Geographic variation in human mitochondrial DNA control region sequence: the population history of Turkey and its relationship to the European populations». *Mol. Biol. Evol.*, 13: 1067-1077.
- DARWIN, C. (1859). *On the Origin of Species*. Londres: John Murray.
- ENARD, W.; PRZEWSKI, M.; FISHER, S. E. [et al.] (2002). «Molecular evolution of *FOXP2*, a gene involved in speech and language». *Nature*, 418: 869-872.
- FALSETTI, A. B.; SOKAL, R. R. (1993). «Genetic structure of human populations in the British Isles». *Ann. Hum. Biol.*, 20: 215-229.
- KARAFET, T. M.; OSIPOVA, L. P.; GUBINA, M. A. [et al.] (2002). «High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life». *Hum. Biol.*, 74: 761-789.
- KRAUSE, J.; LALUEZA-FOX, C.; ORLANDO, L. [et al.] (2007). «The derived *FOXP2* variant of modern humans was shared with Neandertals». *Curr. Biol.*, 17: 1908-1912.
- LAI, C. S.; FISHER, S. E.; HURST, J. A.; VARGHA-KHADEM, F.; MONACO, A. P. (2001). «A forkhead-domain gene is mutated in a severe speech and language disorder». *Nature*, 413: 519-523.
- PÉREZ-LEZAUN, A.; CALAFELL, F.; COMAS, D. [et al.] (1999). «Sex-specific migration patterns in Central Asian populations, revealed by analysis of Y-chromosome short tandem repeats and mtDNA». *Am. J. Hum. Gen.*, 65: 208-219.
- POLONI, E. S.; SEMINO, O.; PASSARINO, G. [et al.] (1997). «Human genetic affinities for Y-chromosome P49a,f/TaqI haplotypes show strong correspondence with linguistics». *Am. J. Hum. Gen.*, 61: 1015-1035.
- RENFREW, C. (1994). «World linguistic diversity». *Sci. Am.*, 270: 116-123.
- ROSSER, Z. H.; ZERJAL, T.; HURLES, M. E. [et al.] (2000). «Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language». *Am. J. Hum. Gen.*, 67: 1526-1543.
- RUBICZ, R.; MELVIN, K. L.; CRAWFORD, M. H. (2002). «Genetic evidence for the phylogenetic relationship between Na-Dene and Yeniseian speakers». *Hum. Biol.*, 74: 743-760.

- RUHLEN, M. (1991). *A guide to the World's languages*, 2a ed. Stanford: Stanford University Press.
- SEIELSTAD, M. T.; MINCH, E.; CAVALLI-SFORZA, L. (1998). «Genetic evidence for a higher female migration rate in humans». *Nat. Gen.*, 20: 278-280.
- SOKAL, R. R.; THOMSON, B. A. (1998). «Spatial genetic structure of human populations in Japan». *Hum. Biol.*, 70: 1-22.
- WOOD, E. T.; STOVER, D. A.; EHRET, C. [et al.] (2005). «Contrasting patterns of Y chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes». *Eur. J. Hum. Gen.*, 13: 867-876.
- ZEI, G.; BARBUJANI, G.; LISA, A. [et al.] (1993). «Barriers to gene flow estimated by surname distribution in Italy». *Ann. Hum. Gen.*, 57: 123-140.